How to use the International Classification of Functioning, Disability and Health as a reference system for comparative evaluation and standardized reporting of rehabilitation interventions

Stucki, Gerold; Pollock, Alex; Engkasan, Julia Patrick; Selb, Melissa

Published in:
European Journal of Physical and Rehabilitation Medicine

DOI:
10.23736/S1973-9087.19.05808-8

Publication date:
2019

Document Version
Author accepted manuscript

How to use the ICF as a reference system for comparative evaluation and standardised reporting of rehabilitation interventions

Gerold STUCKI[1,2,3], Alex POLLOCK[4], Julia Patrick ENGKASAN[5,6], Melissa SELB[2,3]

[1]Department of Health Sciences and Health Policy, University of Luzern (Switzerland)

[2]ICF Research Branch, a cooperation partner within the WHO Collaborating Centre for the Family of International Classifications in Germany (Nottwil, Switzerland)

[3]Swiss Paraplegic Research (Nottwil, Switzerland)

[4]Nursing, Midwifery and Allied Health Professions Research Unit, Glasgow Caledonian University (UK)

[5]Department of Rehabilitation Medicine, University of Malaya (Kuala Lumpur, Malaysia)

University Malaya Medical Centre (Kuala Lumpur, Malaysia)

Corresponding author:

Melissa Selb, Coordinator ICF Research Branch, Guido-Zäch-Strasse 4, 6207 Nottwil, Switzerland, Tel. +41 41 939 66 31, Email: melissa.selb@paraplegie.ch

Abstract

Rehabilitation aims to optimize functioning of persons experiencing functioning limitations. As such the comparative evaluation of rehabilitation interventions relies on the analysis of the differences between the change in patient functioning after a specific rehabilitation intervention versus the change following another intervention. A robust health information reference system that can facilitate the comparative evaluation of changes in functioning in rehabilitation studies and the standardised reporting of rehabilitation interventions is the International Classification of Functioning, Disability and Health (ICF). The objective of this paper is to present recommendations that Cochrane Rehabilitation could adopt for using the ICF in rehabilitation studies by: 1) defining the functioning categories to be included in a rehabilitation study, 2) specifying selected functioning categories and selecting suitable data collection instruments, 3) examining aspects of functioning that have been documented in a study, 4) reporting functioning data collected with various data collection instruments, and 5) communicating results in an accessible, meaningful and easily understandable way. The authors provide examples of concrete studies that underscore these recommendations, whereby also emphasizing the need for future research on the implementation of specific recommendations, e.g. in meta-analysis in systematic literature reviews. Furthermore, the paper outlines how the ICF can complement or be integrated in established Cochrane and rehabilitation research structures and methods, e.g. use of standard mean difference to compare cross-study data collected using different measures, in developing core outcome sets (COS) for rehabilitation, and the use of the PICO model.

Key Words: Functioning, International Classification of Functioning, Disability and Health, rehabilitation, outcomes

Background

Rehabilitation is the health strategy that aims to optimize functioning of persons experiencing or likely to experience a limitation in functioning.[1-3] As an operationalization of the lived experience of health, *functioning* encompasses both a person's intrinsic health capacity and "lived health", i.e. what the person actually does or is limited in doing in light of the interaction between this health capacity and environment factors.[4,5] Complementary to mortality and morbidity, functioning is the third health indicator. At the health systems level, functioning is an indicator of a health system's performance, while at a health interventions level, it is an indicator of impact of interventions and service delivery on a person's health state and corresponding impact on his/her life (with consideration of the person's perspective).[5,6] The importance of functioning in the rehabilitation context has been recognized by the World Health Organization (WHO),[7] Cochrane Rehabilitation[8] and the rehabilitation community at large.[2,3]

Since functioning is a key indicator for rehabilitation,[5,6] the comparative evaluation of the efficacy and effectiveness of diverse rehabilitation interventions relies on the analysis of the differences between the change in patient functioning after completion of a specific rehabilitation intervention and the change following other interventions.

A prerequisite for a meaningful and methodologically sound evaluation of changes in functioning in rehabilitation studies is a robust health information reference system that allows for the documentation of functioning.

Robust health information reference system

A robust health information reference system is essential for the comparison of rehabilitation study data. WHO has developed such a health information reference for functioning data – the *International Classification of Functioning, Disability and Health (ICF)*.[9] Comprising a mutually exclusive and cumulatively exhaustive list of domains (i.e. ICF categories), the ICF provides an internationally accepted conceptual framework for describing a person's health relevant to their lived experience.[3,10] In its initiative *Rehabilitation 2030* WHO emphasized ICF's status as an "internationally standardized language and coding system for data comparability" and its utility in enhancing health information systems.[11]

The ICF not only makes the reporting of both clinical and population health data in an internationally comparative manner possible, the ICF as the operationalization of functioning also allows for the differentiation between a person's intrinsic biological health capacity and lived health.[4,5,12] Most importantly the ICF enables data sharing across a wide range of application areas in research, including clinical trials,[13-16] registries and cohort studies,[17,18] and in developing data collection instruments and methodologies.[18-20]

Despite the universal understanding that functioning information, operationalized through the ICF, is the "core concept underpinning" research in rehabilitation [3], no concrete recommendations on how to use the ICF in rehabilitation research exist to date.

The objective of this paper is to present recommendations on how the ICF can be used as a reference system for comparative evaluation and standardised reporting of interventions in rehabilitation studies that can be adopted by Cochrane Rehabilitation. The specific aims are to provide recommendations on

1) defining the functioning categories to be included in a rehabilitation study
2) specifying selected functioning categories and selecting suitable data collection instruments
3) examining aspects of functioning that have been documented in a study
4) reporting functioning data collected with various data collection instruments
5) communicating results in an accessible, meaningful and easily understandable way

Developed based on the authors' experience of over 15 years of conducting and advising on functioning- and ICF-based research, predominately in and for the rehabilitation context, and in applying Cochrane standards and methodologies, these recommendations are the starting point for establishing comprehensive guidelines and standards on using the ICF in rehabilitation research.

Defining functioning categories to be included in rehabilitation studies

With over 1400 categories across diverse domains of functioning, the ICF provides a comprehensive basis for defining the aspects of functioning to be included in rehabilitation studies as well as in studies that explore health interventions addressing specific functioning outcomes, such as mobility, and self-care. The specification of functioning aspects to document in research is facilitated by ICF-based international standards, including ICF Core

Sets (ICF-CSs) and ICF Generic Sets (www.icf-core-sets.org).[12,21-23]; they provide short lists of ICF categories to be considered for documentation in research and clinical practice. ICF Generic Sets represent minimum standards for the documentation of functioning across clinical and population studies. While the ICF Generic-7 (G7) Set contains the 7 aspects of functioning, i.e. ICF categories, that are ideally collected and documented in population studies.[21], the ICF Generic-30 (G30) Set comprises the 30 categories (which also contains all G7 Set categories) essential to collect and document in clinical studies.[23] (See Table 1). ICF-CSs are a selection of ICF categories from the whole ICF that are considered most relevant for specific health conditions and contexts. ICF-CSs have two versions – the comprehensive and brief version. ICF-CSs are developed by means of a rigorous, multi-method scientific procedure involving four preparatory studies – an empirical multicentre study, a systematic literature review, a qualitative study and an expert survey, and an iterative decision-making and consensus process to decide on the ICF categories to be included in the Comprehensive and Brief ICF-CSs. The Comprehensive ICF-CS is first developed, followed by the brief version, which is a selection of specific categories from the comprehensive version. In terms of rehabilitation research, Brief ICF-CSs provide a minimum standard for studies involving persons with specific health conditions, e.g. stroke or spinal cord injury, or addressing a specific care context, e.g. vocational rehabilitation. Containing more categories, Comprehensive ICF-CSs include additional domains from which to select in order to comprehensively reflect the outcomes being addressed in the study, and if relevant, for the study population.

(Insert Table 1 here)

Table 1. ICF Generic-30 Set with simple, intuitive descriptions

ICF Code	Title	Description
b130	Energy and drive functions (G)	Psychological energy and motivational drive to move towards goals, satisfy needs and control impulses
b134	Sleep functions	Cycle, quality and amount of sleep
b152	Emotional functions (G)	Mental functions for the modulation of the expression of feelings and emotions
b280	Sensation of pain (G)	Unpleasant sensation indicating potential or actual damage of some body structure
b455	Exercise tolerance functions (G)	Capacity of enduring physical exertion related to respiratory and cardiovascular functions
b620	Urination functions	Voluntary control and discharge from the urinary bladder
b640	Sexual functions	Mental and physical functions related to the sexual act
b710	Mobility of joint functions	Range and ease of movement of a joint
b730	Muscle power functions	Capacity to generate force through the contraction of a muscle or muscle groups
d230	Carrying out daily routine (G)	Plan, manage and complete routine daily life activities
d240	Handling stress and other psychological demands	Manage and control the psychological demands to carry out tasks demanding responsibilities involving stress and/or distractions and/or critical issues
d410	Changing basic body position	Changing the body position (for example getting up from a chair, lying down on a bed, kneeling, bending down to pick up an object)
d415	Maintaining a body position	Maintaining a body position in the way and for the time required by the situation
d420	Transferring oneself	Moving from one surface to another while maintaining the same body position
d450	Walking (G)	Moving in an upright position, step by step, always maintaining a support on the ground
d455	Moving around (G)	Moving around differently from walking (for example running, going up and down the stairs, jumping, climbing, swimming, etc.)
d465	Moving around using equipment	Moving around from one place to another, on any surface or space, by using specific tools (skates, skis, or scuba equipment) or assistive devices (wheelchair etc.)
d470	Using transportation	Using different means of transportation to move around as a passenger
d510	Washing oneself	Cleaning, washing and drying one's whole body, or body parts
d520	Caring for body parts	Caring for skin, teeth, hair, finger and toe nails, genitals, etc. that require more than washing and drying
d530	Toileting	Management of urination, defecation and menstruation including cleaning oneself afterwards
d540	Dressing	Choosing, putting on and taking off clothes and footwear in accordance with climatic and social conditions
d550	Eating	Eating food already served, in a coordinated sequence and adapted to the context
d570	Looking after one’s health	Ensuring health and physical and mental well-being by adopting a healthy lifestyle
d640	Doing housework	Managing a household by cleaning and clearing up, washing clothes, using household appliances, disposing of garbage, etc.

d660	Assisting others	Assisting others with their learning, communicating, self-care, movements and being concerned about their well-being
d710	Basic interpersonal interactions	Interacting with people in a contextually and socially appropriate manner
d770	Intimate relationships	Creating and maintaining close or romantic relationships between individuals, such as husband and wife, lovers or sexual partners
d850	Remunerative employment (G)	Properly performing remunerative employment (full or part time or self-employed) in all its aspects
d920	Recreation and leisure	Engaging in recreational or leisure activity (play, cultural and sports activities etc. during spare time

G = ICF Generic-7 Set category

As ICF sets are considered the international standard for documenting and reporting functioning, and use in rehabilitation practice along the continuum of care,[3,10] it would make sense that the categories contained in selected ICF sets are used to define specific functioning outcomes in rehabilitation and rehabilitation-relevant studies.

Population study examples include the Swiss Spinal Cord Injury (SCI) Cohort Study (SwiSCI) and the International SCI Community Survey (InSCI). In SwiSCI the ICF G30 Set and the Brief ICF-CS for SCI in early post-acute and long-term contexts were used to determine which variables to include in the study.[17,24] A similar procedure was employed in developing the questionnaire used in InSCI, a multi-centre study to collect internationally comparable SCI data all over the world.[18] However, since InSCI is a community survey, only the G30 Set and ICF-CS for SCI long-term context were employed. Although both SwiSCI and InSCI surveys do not display ICF codes, the codes are embedded in the data management system for possible use in comparing the data collected with general population and other SCI population data.

With regard to clinical studies in rehabilitation, an illustration of how to define the functioning categories is provided by an analysis of physical therapy and occupational therapy delivered as part of a randomised controlled trial (RCT).[25] Therapists delivering "conventional" rehabilitation for people with arm impairment following stroke recorded the duration of each therapy session spent focussed on specific ICF categories. Prior to the trial, the therapists in the participating rehabilitation centres reached consensus on a therapy recording schedule which comprised fourteen ICF categories. This included ICF body function components (e.g. b1 mental functions, b2 sensory functions and pain), activities components (e.g. d4 mobility, d5 self-care) and participation components (d9 community, social and civil life).[25]

There may already be ICF-CS based instruments that can be used to define the functioning outcomes in rehabilitation studies. An example of this is the multi-centre RCT conducted by McNeely and colleagues that examined the efficacy of night-time compression on arm lymphedema volume maintenance and quality of life (QoL) in breast cancer survivors. In this study, the investigators employed the Lymph-ICF to assess the primary outcome of QoL. Lymph-ICF was developed based on the ICF set for lymphedema to assess QoL and to monitor progress of treatment addressing lymphedema-related functioning (e.g. mobility, pain).[15]

ICF sets have also been shown to be useful in defining functioning outcomes beyond rehabilitation studies. For example, in a multicentre cross-sectional study a questionnaire containing categories of the ICF-CS for head and neck cancer was used as one of the outcome instruments to compare functional outcome in patients with advanced head and neck cancer treated with a) surgical resection and reconstruction with microvascular free flaps followed by radiochemotherapy versus b) primary radiochemotherapy. The study investigators concluded that the ICF offers a multidimensional view that can guide rehabilitation according to patients' needs.[16]

These examples show that when defining functioning outcomes in rehabilitation and other healthcare studies it is recommended that:

- ICF sets (ICF Generic Sets and Core Sets) are considered for defining the categories
- the selection of the appropriate ICF set(s) depends on the study population and context
- the corresponding ICF categories are preferably defined using ICF codes explicitly or as analysable information in the data management system

If a rehabilitation study is focusing only on one specific aspect of functioning, it may be reasonable to consider only the corresponding ICF category(ies) rather than a whole ICF set. For example, in a systematic literature review and meta-analysis of RCTs on the effect of diverse physical therapy (PT) interventions on balance of patients with multiple sclerosis (MS), Paltamaa and colleagues[26] operationalized balance with the ICF categories d410-d429 Changing and maintaining body position. Although not explicit in the publication on the review, these categories are contained in the ICF-CS for MS. Accordingly, the investigators included in the review only RCTs that employed outcome measures that were linked to these categories. These categories not only facilitated the selection process, they helped to assure the comparability of the RCTs as the included RCTs were all linked to d410-d429.

Specifying defined functioning categories and selecting suitable data collection instruments

ICF-CSs provide the starting point for what to document and report; however, they do not prescribe how to measure the aspects of functioning selected.[27] Once the ICF categories that reflect the study's functioning outcomes have been defined, suitable instruments for collecting the functioning outcomes data need to be selected. When measuring aspects of functioning a challenge arises in selecting suitable

data collection instruments for a specific study from a range of currently available and sound clinical tests and patient-reported outcome measures. It is important to recognize that there is not one data collection instrument that fits all purposes. There may be several instruments that are suitable for collecting data for a specific purpose while some instruments can be used in diverse areas. Moreover, there is a continuous evolution of instruments – while some instruments become "legacy instruments" (i.e. commonly and widely used instruments, e.g. the 36-item Short Form Survey called SF-36)[28], new instruments are developed. The launch of ICF has triggered the development of ICF-based instruments, such as the ICF-CS based instrument developed by the Assessment of SpondyloArthritis international Society (ASAS) called ASAS Health Index.[20] This instrument has been validated and can be used in clinical trials, e.g.to evaluate the impact of different treatments on the functioning of patients with spondyloarthritis.[29]

In selecting the instrument(s) it is recommended that:

- potential data collection instruments are mapped (or linked) to the study's set of ICF categories
- an efficient battery of data collection instruments is selected, i.e. preferably only one clinical test or questionnaire item per category

Mapping (referred to as "linking" from now on) instruments to the ICF at any level of granularity, i.e. item, domain, chapter or component level), identifies the aspects of functioning that is covered by the instruments. Linking involves identifying the meaningful concept(s) in the item of the instrument and deciding which ICF category(ies) that corresponds to the meaningful concept(s) by using established ICF linking rules.[30] For example, the item "lifting and carrying groceries" in SF-36[28] contains the main concepts of lifting and carrying objects. These concepts correspond to the ICF category d430 Lifting and carrying objects. Ideally, the items are linked independently by two persons, the individual linking results are compared and a decision is made about the final category(ies) for each item. A third person is consulted if necessary to find consensus on the final category(ies). Knowing the functioning aspects covered by an instrument helps to ensure that the instrument(s) planned for the collection of study data cover all the ICF categories defined for the study, while also avoiding the redundant coverage of the same categories. For example, in a study that evaluated the effect of an intensive hybrid model of aphasia therapy on the communication and well-being in a hospital setting, the investigators

administered the Comprehensive Aphasia Test and the Boston Naming Test, since these specifically assessed the ICF categories the study was targeting, i.e. 'reception of language' and 'expression of language'.[31]

Focusing only on instruments that cover the specific functioning aspects being addressed in a study leads to a more efficient battery of instruments that may, in turn, reduce participant burden.

However, for specific aspects of functioning, a plethora of data collection instruments may make identifying targeted instruments to employ in trials difficult. This, in turn, has implications for data analysis. This is illustrated by the wide variety of diverse data collection instruments that have been used to record aspects of arm function in stroke studies. Systematic reviews on this topic identified between 48[32] and 144[33] different measures. The uncoordinated use of a high number of different instruments limit the ability to pool data from trials in a meaningful way, and the need for targeted selection based on a narrower collection of measures has been highlighted[33]. Linking the instruments to ICF provides an efficient and effective way of supporting the selection of meaningful measures for trials. For example, Santisteban and colleagues[32] found that 11 of the 15 most commonly used measures only evaluated the ICF activity domains of d4 Mobility, d5 Self-care, d6 Domestic life and/or b7 Neuromusculoskeletal movement-related functions. Moreover, in studies such as the aforementioned literature review examining the effect of PT interventions on balance of MS patients,[26] ICF linking of the instruments included in the screened trials can facilitate the process of selecting which trials to include or exclude in the review.

A rehabilitation study may also call for developing a tailored ICF-based data collection tool or for employing only a limited set of items/questions that reflect selected ICF categories rather than a whole instrument. To ensure that the item/question adequately and soundly addresses what is intended, it is recommended that each ICF category is specified by way of 'category specification', a method by which items from existing instruments are used to operationalize the ICF categories selected.[34] For example, b4105 Bending can be operationalized with the item 'Bend down to pick up clothing from the floor' from the Lequesne Index of Severity for Osteoarthritis of the Hip if the study intends to assess bending ability. Category specification involves performing a literature search to identify measures used to assess variables investigated in previous studies that correspond to the selected categories. To

ascertain if and which items of the measures truly reflect the selected categories, the measures are linked to the ICF.[30] Expert input can strengthen the decision which item/question(s) best operationalizes the selected categories. Figure 1 shows how category specification can be done using the category b4015 Bending as an example.

(Insert figure 1 here)

It is important to note that the psychometric properties of the measures from which the items are derived cannot just be taken over into the new instrument. The new instrument still has to be tested for psychometric properties.

Category specification is exemplified by the development of the surveys used in SwiSCI [17] and INSCI.[18] In INSCI, for example, an item from the Spinal Cord Injury Secondary Scale [35] and an item from the Brief Pain Inventory [36] were used to operationalize the two INSCI survey questions that covered the category b280 Sensation of pain. The scales for operationalizing the ICF Core Set categories used as the variables for developing the INSCI survey questions were selected in a two-step process. The first step involved a pre-selection of scales that were known to the InSCI research team who had expertise in epidemiology and SCI. The selection of scales were guided by the following criteria: 1) efficacy (number of questions required to cover a particular category) versus granularity, 2) feasibility (availability in the participating country languages, question complexity), 3) comparability (use of legacy scales versus less known scales, SCI-specific scales versus more generic scales), and 4) truth and discrimination (meet standards for validity and reliability). The InSCI research team presented and discussed this pre-selection, including information on the aforementioned criteria for each selected scale, with the national leaders and coordinators of the participating InSCI countries at two international conferences. The final selection of scales was made based on the feedback obtained during these conferences.[18]

Examining functioning information documented in a study

There may be studies in which ICF categories to include in a study are not pre-define but rather the researchers decide to openly examine the aspects of functioning that have been impacted by one or more interventions. In this case, we recommend that:

- these functioning aspects are linked to the ICF for enhanced comparative evaluation and also for reporting (see below)
- the comparative evaluation includes a comparison of the resulting ICF categories with established ICF sets indicated above

For example, there are two Cochrane Systematic Review reviews investigating rehabilitation interventions to improve walking or mobility in persons after stroke. English and colleagues [37] measured walking distance, speed and balance, whilst Mehrholz and colleagues [38] measured independence in walking, and walking speed and distance. Even though both groups did not specify ICF categories for functioning they intended to measure at the conception stage, the outcomes measured in these studies can be linked to ICF categories d450 walking or more specifically d4500 Walking short distances.

Implementing the recommendations described above for defining functioning categories to include in studies, specifying them, selecting appropriate data collection instruments, and examining functioning information documented in studies would optimise the comparative evaluation of functioning changes and interventions in rehabilitation studies. Crucial to ensuring the utility of the results in rehabilitation studies is the approach taken for report and communicating the results.

Reporting of functioning data collected with various data collection instruments

When reporting the functioning outcomes addressed in rehabilitation studies it is recommended that:

- the functioning data collected and used in the analysis are clearly shown, preferably identified with ICF categories
- results are presented using the scoring systems of the respective original data collection instruments as well as using the ICF-based common metric generated according to the ICF Standardized Assessment and Reporting System (ICF-StARS).

Reporting of functioning data collected actually begins with the study protocol, i.e. planning for what to report. The aforementioned illustrative examples of studies highlight the importance of clearly

identifying the pre-defined functioning outcomes of interest as ICF categories and the measures that will be used to assess these outcomes.

Clearly identifying the ICF categories of the functioning data collected and examined in studies would make it easier for other researchers to find functioning-related studies that are relevant for planning and conducting their own studies, and for discussing results. Moreover, with an increasing number and range of studies published in rehabilitation-relevant scientific journals and in other peer-review healthcare journals that apply the ICF, [39,40] adapting existing trial repositories/registries, such as www.clinicaltrials.gov, to include the ICF as a study identifier may be worthy of consideration. In terms of rehabilitation studies, clearly identifying the categories applied in studies may prove useful should a repository of rehabilitation studies organized according to functioning aspects be developed. A few databases that group outcome measures in rehabilitation already exist, e.g. https://www.strokengine.ca/en/find-assessment/, https://www.sralab.org/rehabilitation-measures. However, none of these databases currently group measures according to the ICF.

Given that data on functioning may be collected across studies using a range of different data collection instruments (legacy or otherwise), it is essential that the results are reported in such a way that it allows for the comparison of these cross-study data in meta-analyses. To be able to compare data on functioning across studies that have been collected using different instruments, one approach would be to transform the data (operationalized with the ICF) to common ICF-based metrics. This is possible through use of ICF-StARS. ICF-StARS has been developed, tested and used by the ICF Research Branch – a cooperation partner within the WHO Collaborating Centre for the Family of International Classifications in Germany (hosted by the German Institute for Medical Documentation and Information, abbreviated in German as DIMDI). It involves linking data collected with any instrument to ICF categories (alternatively to broader domains, like self-care or mobility) to identify concept equivalence, i.e. the items of the instruments that are being compared measure the same construct (categories or domains). Subsequently, the scores of conceptually equivalent (sub)scales of the different instruments are analysed using the Rasch model[41] to achieve scale equivalence between the instruments, i.e. a score in one instrument is equivalent to the score of the same domain in another instrument. This procedure results in a table that transforms the original raw scores of the (sub)scales, generally ordinal-based, to an interval-scaled reference metric score. Until now the procedure for achieving score

equivalence using Rasch has been done only on total scale and subscale scores. Additional research would be needed to bring clarity on score equivalence at the individual item level. The reference metric score of each instrument can then be compared.[12,19]

Another approach to comparing and combining data from studies that have been collected using different measures is the standard approach employed in Cochrane reviews of interventions, which involves calculation of standard mean differences (SMD). [42] The SMD is a summary statistic within meta-analyses, and involves standardising the results of studies to a uniform scale before they are combined. The SMD expresses the size of the treatment effect in a study relative to the variability observed in that study, i.e. the difference in mean outcomes between intervention and comparator (placebo) intervention relative to the standard deviation observed among participants.

ICF-StARS can be considered complementary to the SMD approach, providing the interval scale scores as the basis for the mean outcomes for calculating the SMD. Due to its mathematical measurement properties (linearity and equidistant scale thresholds), scores derived from an interval scale are preferred over ordinal-scales scores. More importantly, interval scaled scores enable parametric analysis that is more likely to detect significant differences or changes in scores.[41]

Both these approaches can strengthen the utility of clinical trial data collected with different instruments, specifically the evaluation of changes in functioning, by enabling their comparability.

Communicating results meaningfully

Lastly, an important element to enhance the utility of clinical trial data is the communication of results in a meaningful, understandable and easily accessible manner for professionals, scientists and consumers alike. The presentation of study results using ICF categories and domains can be a conduit for communicating what has been addressed in a study. The ICF is universally accepted in the rehabilitation community, and ICF categories and domains are described so that it can also be understood by lay experts. Moreover, presenting functioning on a common metric, as described above, offers the possibility of communicating functioning levels of study populations and changes in functioning on a universally understandable 0-100 scale, while respecting the value placed on specific instruments, especially legacy instruments, for collecting data.[17]

Although clinical trial data are often communicated in terms of clinical meaningful/important difference or effect size, this may be challenging when we attempt to present data on specific aspects of functioning, such as those related to activities and participation, in terms of clinical meaningful/important difference or effect size. Indeed, there are particular challenges to clinical interpretation of SMDs calculated within meta-analyses pooling data from studies using different measures, as the calculated effect is reported in units of standard deviation, rather than in units of a specific scale of clinical measurement. There is also increasing recognition that providing quality evidence for clinical practice extends to evaluating impact of interventions on patients' well-being [6,8] – "how they live their health in their actual environment". [6] Cochrane Rehabilitation, who aims to use rehabilitation research to bridge what is clinically important and what is important for patients' well-being, is an ideal platform for innovative methodological discussions and future research on this topic, specifically on how to use an ICF-based common metric to translating data into clinically important changes.

Discussion

With the understanding that functioning is a key indicator for rehabilitation's performance and impact on a person's health state and lived experience of health [5,6] and that the ICF is the health information reference system for documenting functioning, the rehabilitation community is spearheading the application of the ICF in the comparative evaluation and standardised reporting of interventions in rehabilitation research.

In this paper we have argued that the ICF is a robust health information reference system in documenting functioning in rehabilitation studies. We illustrated and provided concrete recommendations on how to use the ICF in defining the functioning categories to be included in rehabilitation studies, in specifying these defined functioning categories, in selecting suitable data collection instruments, and in reporting study results.

Relevance in medicine beyond rehabilitation

Since functioning information plays an important role in other medical fields beyond rehabilitation,[5,14] these recommendations and the overall approach of ICF implementation in clinical research presented

in this paper have potential value for clinical research beyond the area of rehabilitation. In the context of Cochrane, this information may be of value not only for Cochrane Rehabilitation (https://rehabilitation.cochrane.org/)[43] but also for diverse Cochrane fields, specifically for consideration by diverse methodology groups. This notion is especially relevant considering the growing use of patient-reported outcomes measures (PROMs) in healthcare research in general.[44,45] Patient-reported outcomes (PROs), often operationalised as functioning information ('functional status' or 'functional ability') are considered important as they provide valuable insight in the patient's perspective of health and reflections on clinical interventions. Cochrane has a methods group "PRO Methods Group" (http://methods.cochrane.org/pro/welcome), whose mission is to "incorporate self-reported measures from patients to help determine whether treatments are doing more good than harm". The recommendations presented in this paper are consistent with the principles espoused by the PRO Methods Group; specifically it encourages the specification and labelling of content and type of measure for every application of a PRO when conducting meta-analyses.

Link to PICO

Another aspect of research methodology in which the ICF may find utility is in the PICO (Patient, Population or Problem/Intervention/Comparison/Outcome) model. PICO is a strategy for wording research questions and developing search strategies as well as characterising clinical trial or meta-analyses (http://community.cochrane.org/tools/data-management-tools/pico-annotation-project). Functioning information, operationalised through the ICF, can be employed for example to describe the patient population, components of some interventions, and outcomes from a functioning perspective. Concretely, the ICF can be added to the Cochrane PICO Finder (http://linkeddata.cochrane.org/picofinder) under the rubric "Intervention/Comparator" as well as under "Outcomes".

ICF and Core Outcome Sets (COS)

Consistent selection of appropriate outcomes in clinical trials and systematic reviews are essential if study findings are to be useful, reliable, and relevant to key stakeholders, such as patients, healthcare professionals, and healthcare decision-makers. Accordingly, core outcome sets (COS), i.e. agreed minimum standard set of outcomes to be measured and reported in clinical trials in specific areas of health, have increasingly captured attention of researchers, including in rehabilitation.[46-50] The

recommendations introduced in this paper are in line with the sentiment behind COS development, i.e. to facilitate comparability, ultimately improving the standards of reporting of studies. As such, the ICF could play a central role in future COS development projects and updating existing COS, especially those related to functioning.

The value of using the ICF to specify outcomes and define concepts in studies was recognized early on by Outcome Measures in Rheumatology (OMERACT), recommending in 2009 that the ICF is employed when defining functioning and disability in emerging OMERACT core domain sets.[51-53] OMERACT continues to show its support for using the ICF in developing its core domain set for clinical trials of shoulder disorders – proposals for aligning the definition of participation in this core domain set were adopted.[48]

An example of the key role the ICF can play in COS development is the project of Wallace and colleagues, who used the ICF to identify outcomes prioritized by diverse stakeholder groups in aphasia treatment. They found that the stakeholder groups prioritized for example b152 Emotional functions, various categories under d3 Communication, and d580 Health services, systems, and policies as outcomes to measure when assessing the impact of aphasia treatment.[50] COS projects can be found in databases such as hosted by (Core Outcome Measures in Effectiveness Trials) http://www.cometinitiative.org/.

Conclusion

This is the first paper that comprehensively and concretely outlines recommendations that Cochrane Rehabilitation could consider for developing possible guidelines on applying the ICF as a reference system for comparative evaluation and standardised reporting of interventions in rehabilitation studies. Introducing more detailed methodological notes on individual recommendations may be warranted. Moreover, future research would provide empirical evidence for implementing specific recommendations in diverse areas of application, e.g. employing the ICF-StARS methodology in meta-analysis in Cochrane Reviews.

Moreover, these recommendations are the beginning of a continuous learning process toward improving the application of the ICF in rehabilitation research. To help enhance these and other recommendations presented in this paper, we encourage readers with expertise as researchers and clinicians in

rehabilitation and other disciplines to submit letters to the *European Journal of Physical and Rehabilitation Medicine* or Cochrane Rehabilitation with comments on the content of this paper and possible suggestions for further developing these recommendations.

Conflicts of interest: The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Acknowledgements: This paper came to fruition with the support of the Royal Society Te Apārangi, New Zealand, who provided grant funding for the Cochrane Catalyst Workshop held in July 2018 in Paris (France), at which the general outline of this paper was presented. The authors also thank the participants of the Cochrane Catalyst Workshop for their comments which helped to shape the outline of this paper, as well as the peer reviewers who provided valuable feedback toward refining the paper as a whole. Finally, the authors would like to thank Carolina Fellinghauer, Núria Duran Androher and Cristina Ehrmann for their input on the standardized reporting content of the paper.

References

1. Stucki G, Cieza A, Melvin J. The International Classification of Functioning, Disability and Health (ICF): a unifying model for the conceptual description of the rehabilitation strategy. *J Rehabil Med.* 2007;39(4):279-285.
2. Meyer T, Gutenbrunner C, Bickenbach J, Cieza A, Melvin J, Stucki G. Towards a conceptual description of rehabilitation as a health strategy. *J Rehabil Med.* 2011;43(9):765-769.
3. European Physical and Rehabilitation Medicine Bodies Alliance. White Book on Physical and Rehabilitation Medicine (PRM) in Europe. *Eur J Phys Rehabil Med.* 2018;54(2):125-321.
4. Bostan C, Oberhauser C, Stucki G, Bickenbach J, Cieza A. Biological health or lived health: which predicts self-reported general health better? *BMC Public Health.* 2014;14:189.
5. Stucki G, Bickenbach J. Functioning: the third health indicator in the health system and the key indicator for rehabilitation. *Eur J Phys Rehabil Med.* 2017;53(1):134-138.
6. Stucki G, Bickenbach J. Health, Functioning and Wellbeing: Individual and Societal. *Arch Phys Med Rehabil.* 2019. In print.
7. Krug E, Cieza A. Strengthening health systems to provide rehabilitation services. *Bull World Health Organ.* 2017;95(3):167.
8. Negrini S, Arienti C, Gimigliano F, Grubisic F, Howe T, Ilieva E, Levack W, Malmivaara A, Meyer T, Patrick Engkasan J, Rathore FA, Kiekens C. Cochrane Rehabilitation: Organization and Functioning. *Am J Phys Med Rehabil.* 2018;97(1):68-71.
9. World Health Organization. *International Classification of Functioning, Disability and Health*. Geneva: World Health Organization; 2001.
10. Imamura M, Gutenbrunner C, Stucki G, Li J, Lains J, Frontera W, Olver J, Ozcakar L, DeLisa J, Battistella LR, Melvin J. The International Society of Physical and Rehabilitation Medicine: the way forward - II. *J Rehabil Med.* 2014;46(2):97-107.
11. World Health Organization. Health information systems and rehabilitation. *Rehabilitation 2030: a call for action.* Geneva, Switzerland. 2017.
12. Stucki G, Prodinger B, Bickenbach J. Four steps to follow when documenting functioning with the International Classification of Functioning, Disability and Health. *Eur J Phys Rehabil Med.* 2017;53(1):144-149.
13. Burke D, Gorman E, Stokes D, Lennon O. An evaluation of neuromuscular electrical stimulation in critical care using the ICF framework: a systematic review and meta-analysis. *Clin Respir J.* 2016;10(4):407-420.
14. Idzerda L, Rader T, Tugwell P, Boers M. Can we decide which outcomes should be measured in every clinical trial? A scoping review of the existing conceptual frameworks and processes to develop core outcome sets. *J Rheumatol.* 2014;41(5):986-993.
15. McNeely ML, Campbell KL, Webster M, Kuusk U, Tracey K, Mackey J. Efficacy of night-time compression for breast cancer related lymphedema (LYNC): protocol for a multi-centre, randomized controlled efficacy trial. *BMC Cancer.* 2016;16:601.
16. Tschiesner U, Schuster L, Strieth S, Harreus U. Functional outcome in patients with advanced head and neck cancer: surgery and reconstruction with free flaps versus primary radiochemotherapy. *Eur Arch Otorhinolaryngol.* 2012;269(2):629-638.

17. Prodinger B, Ballert CS, Cieza A. Setting up a cohort study of functioning: From classification to measurement. *J Rehabil Med.* 2016;48(2):131-140.

18. Fekete C, Post MW, Bickenbach J, Middleton J, Prodinger B, Selb M, Stucki G. A Structured Approach to Capture the Lived Experience of Spinal Cord Injury: Data Model and Questionnaire of the International Spinal Cord Injury Community Survey. *Am J Phys Med Rehabil.* 2017;96(2 Suppl 1):S5-S16.

19. Prodinger B, Tennant A, Stucki G. Standardized reporting of functioning information on ICF-based common metrics. *Eur J Phys Rehabil Med.* 2018;54(1):110-117.

20. Kiltz U, van der Heijde D, Boonen A, Cieza A, Stucki G, Khan MA, Maksymowych WP, Marzo-Ortega H, Reveille J, Stebbings S, Bostan C, Braun J. Development of a health index in patients with ankylosing spondylitis (ASAS HI): final result of a global initiative based on the ICF guided by ASAS. *Ann Rheum Dis.* 2015;74(5):830-835.

21. Cieza A, Oberhauser C, Bickenbach J, Chatterji S, Stucki G. Towards a minimal generic set of domains of functioning and health. *BMC Public Health.* 2014;14:218.

22. Selb M, Escorpizo R, Kostanjsek N, Stucki G, Ustun B, Cieza A. A guide on how to develop an International Classification of Functioning, Disability and Health Core Set. *Eur J Phys Rehabil Med.* 2015;51(1):105-117.

23. Prodinger B, Cieza A, Oberhauser C, Bickenbach J, Ustun TB, Chatterji S, Stucki G. Toward the International Classification of Functioning, Disability and Health (ICF) Rehabilitation Set: A Minimal Generic Set of Domains for Rehabilitation as a Health Strategy. *Arch Phys Med Rehabil.* 2016;97(6):875-884.

24. Prodinger B, Ballert CS, Brach M, Brinkhof MW, Cieza A, Hug K, Jordan X, Post MW, Scheel-Sailer A, Schubert M, Tennant A, Stucki G. Toward standardized reporting for a cohort study on functioning: The Swiss Spinal Cord Injury Cohort Study. *J Rehabil Med.* 2016;48(2):189-196.

25. de Jong LD, van Wijck F, Stewart RE, Geurts ACH, Dijkstra PU. Content of conventional therapy for the severely affected arm during subacute rehabilitation after stroke: An analysis of physiotherapy and occupational therapy practice. *Physiother Res Int.* 2018;23(1).

26. Paltamaa J, Sjogren T, Peurala SH, Heinonen A. Effects of physiotherapy interventions on balance in multiple sclerosis: a systematic review and meta-analysis of randomized controlled trials. *J Rehabil Med.* 2012;44(10):811-823.

27. Stucki G. International Classification of Functioning, Disability, and Health (ICF): a promising framework and classification for rehabilitation medicine. *Am J Phys Med Rehabil.* 2005;84(10):733-740.

28. Ware JE, Jr. SF-36 health survey update. *Spine (Phila Pa 1976).* 2000;25(24):3130-3139.

29. Kiltz U, van der Heijde D, Boonen A, Akkoc N, Bautista-Molano W, Burgos-Vargas R, Wei JC, Chiowchanwisawakit P, Dougados M, Duruoz MT, Elzorkany BK, Gaydukova I, Gensler LS, Gilio M, Grazio S, Gu J, Inman RD, Kim TJ, Navarro-Compan V, Marzo-Ortega H, Ozgocmen S, Pimentel Dos Santos F, Schirmer M, Stebbings S, Van den Bosch FE, van Tubergen A, Braun J. Measurement properties of the ASAS Health Index: results of a global study in patients with axial and peripheral spondyloarthritis. *Ann Rheum Dis.* 2018;77(9):1311-1317.

30. Cieza A, Fayed N, Bickenbach J, Prodinger B. Refinements of the ICF Linking Rules to strengthen their potential for establishing comparability of health information. *Disabil Rehabil.* 2019;41(5):574-583.

31. Wenke R, Cardell E, Lawrie M, Gunning D. Communication and well-being outcomes of a hybrid service delivery model of intensive impairment-based treatment for aphasia in the hospital setting: a pilot study. *Disabil Rehabil.* 2018;40(13):1532-1541.

32. Santisteban L, Teremetz M, Bleton JP, Baron JC, Maier MA, Lindberg PG. Upper Limb Outcome Measures Used in Stroke Rehabilitation Studies: A Systematic Literature Review. *PLoS One.* 2016;11(5):e0154792.

33. Duncan Millar J, van Wijck F, Pollock A, Ali M. Outcome measures in post-stroke arm rehabilitation trials: do existing measures capture outcomes that are important to stroke survivors, carers, and clinicians? *Clin Rehabil.* 2019:269215518823248.

34. Cieza A, Boldt C, Ballert CS, Eriks-Hoogland I, Bickenbach JE, Stucki G. Setting up a cohort study on functioning: deciding what to measure. *Am J Phys Med Rehabil.* 2011;90(11 Suppl 2):S17-28.

35. Kalpakjian CZ, Scelza WM, Forchheimer MB, Toussaint LL. Preliminary reliability and validity of a Spinal Cord Injury Secondary Conditions Scale. *J Spinal Cord Med.* 2007;30(2):131-139.

36. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med Singapore.* 1994;23(2):129-138.

37. English C, Hillier SL, Lynch EA. Circuit class therapy for improving mobility after stroke. *Cochrane Database Syst Rev.* 2017;6:CD007513.

38. Mehrholz J, Thomas S, Elsner B. Treadmill training and body weight support for walking after stroke. *Cochrane Database Syst Rev.* 2017;8:CD002840.

39. Cerniauskaite M, Quintas R, Boldt C, Raggi A, Cieza A, Bickenbach JE, Leonardi M. Systematic literature review on ICF from 2001 to 2009: its use, implementation and operationalisation. *Disabil Rehabil.* 2011;33(4):281-309.

40. Madden RH, Bundy A. The ICF has made a difference to functioning and disability measurement and statistics. *Disabil Rehabil.* 2018:1-13.

41. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum.* 2007;57(8):1358-1362.

42. Higgins J, Green SE. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. 2011; https://handbook-5-1.cochrane.org/chapter_9/9_2_3_2_the_standardized_mean_difference.htm.

43. European Physical and Rehabilitation Medicine Bodies Alliance. White Book on Physical and Rehabilitation Medicine (PRM) in Europe. Chapter 10. Science and research in PRM: specificities and challenges. *Eur J Phys Rehabil Med.* 2018;54(2):287-310.

44. Johnston BC, Patrick DL, Busse JW, Schunemann HJ, Agarwal A, Guyatt GH. Patient-reported outcomes in meta-analyses--Part 1: assessing risk of bias and combining outcomes. *Health Qual Life Outcomes.* 2013;11:109.

45. Vodicka E, Kim K, Devine EB, Gnanasakthy A, Scoggins JF, Patrick DL. Inclusion of patient-reported outcome measures in registered clinical trials: Evidence from ClinicalTrials.gov (2007-2013). *Contemp Clin Trials.* 2015;43:1-9.

46. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, Williamson PR. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLoS Med.* 2017;14(11):e1002447.

47. Connolly B, Denehy L, Hart N, Pattison N, Williamson P, Blackwood B. Physical Rehabilitation Core Outcomes In Critical illness (PRACTICE): protocol for development of a core outcome set. *Trials.* 2018;19(1):294.

48. Ramiro S, Page MJ, Whittle SL, Huang H, Verhagen AP, Beaton DE, Richards P, Voshaar M, Shea B, van der Windt D, Kopkow C, Lenza M, Jain NB, Richards B, Hill C, Gill TK, Koes B, Foster NE, Conaghan PG, Smith T, Malliaras P, Roe Y, Gagnier JJ, Buchbinder R. The OMERACT Core Domain Set for Clinical Trials of Shoulder Disorders. *J Rheumatol.* 2019. In print

49. Smith TO, Hawker GA, Hunter DJ, March LM, Boers M, Shea BJ, Christensen R, Guillemin F, Terwee CB, Williamson PR, Dodd S, Roos EM, Loeser RF, Schnitzer TJ, Kloppenburg M, Neogi T, Ladel CH, Kalsi G, Kaiser U, Buttel TW, Ashford AE, Mobasheri A, Arden NK, Tennant A, Hochberg MC, de Wit M, Tugwell P, Conaghan PG. The OMERACT-OARSI Core Domain Set for Measurement in Clinical Trials of Hip and/or Knee Osteoarthritis. *J Rheumatol.* 2019. In print.

50. Wallace SJ, Worrall L, Rose T, Le Dorze G. Using the International Classification of Functioning, Disability, and Health to identify outcome domains for a core outcome set for aphasia: a comparison of stakeholder perspectives. *Disabil Rehabil.* 2017:1-10.

51. Stucki G, Boonen A, Tugwell P, Cieza A, Boers M. The World Health Organisation International Classification of Functioning, Disability and Health: a conceptual model and interface for the OMERACT process. *J Rheumatol.* 2007;34(3):600-606.

52. Boonen A, Stucki G, Maksymowych W, Rat AC, Escorpizo R, Boers M. The OMERACT-ICF Reference Group: integrating the ICF into the OMERACT process: opportunities and challenges. *J Rheumatol.* 2009;36(9):2057-2060.

53. Boers M, Idzerda L, Kirwan JR, Beaton D, Escorpizo R, Boonen A, Magasi S, Sinha I, Stucki G, Tugwell P. Toward a generalized framework of core measurement areas in clinical trials: a position paper for OMERACT 11. *J Rheumatol.* 2014;41(5):978-985.